

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Om protein - protein search, using sw model

Run on:

March 1, 2001, 15:52:31 ; Search time 170.72 Seconds

(without alignments)

25.057 Million cell updates/sec

Title: US-09-331-631a-3_COPY_186_248

Perfect score: 353
Sequence: KRDPQREYEDRRRCQEQE. LINPQRGSGRVEEGERKQS 63

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 195891 seqs, 6790655 residues

Total number of hits satisfying chosen parameters: 195891

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%, Maximum Match 100%, Listing first 45 summaries

Database : PIR_66:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description	RESULT	1	ALIGNMENTS
1	118.5	33.6	566	2	S22477	S22477		
2	118	33.4	509	2	S08059	vicilin precursor	- cacao	
3	110	31.2	588	1	FNCNAB	C-Species: Theobroma cacao (cacao)		
4	103	29.2	637	2	S25221	C-Accession: S22477; S22478; S18105; S22050		
5	101.5	28.8	605	2	S06398	R-McHenry, L.; Fritz, P.J.		
6	89.5	25.4	810	2	T04430	Plant Mol. Biol. 18, 1173-1176, 1992		
7	84	23.8	236	2	T01662	A-title: Comparison of the structure and nucleotide sequence of vicilin genes of caco		
8	81	22.9	966	2	S25365	A-Reference number: S22477; MUID:92288309		
9	78.5	22.2	165	2	T03998	A-Accession: S22477		
10	78	22.1	411	2	T02475	A-Molecule type: DNA		
11	78	22.1	540	2	S1825	A-Accession: S22477		
12	78	22.1	573	2	A53234	A-Residues: 1-666 <MCH>		
13	75	21.2	562	2	T02642	A-Cross-references: EMBL:X62625		
14	75	21.2	568	2	T026243	A-Accession: 1-452 <MC2>		
15	74.5	21.1	407	2	T02258	A-Cross-references: EMBL:X62626		
16	74.5	21.1	582	2	B53234	C-Accession: S22478		
17	74	21.1	686	2	T19371	C-Genetics: 211/1, 269/3; 296/3; 391/3; 502/1		
18	74	21.0	604	2	T05132	C-Introns: 211/1, 269/3; 296/3; 391/3; 502/1		
19	74	21.0	707	2	A08866	C-Keywords: seed; storage protein		
20	74	21.0	1090	2	A41696	F-1-24/Domain: signal sequence #status predicted <SIG>		
21	74	21.0	1671	2	T01628	F-25-566/Product: vicilin #status predicted <MAT>		
22	73	20.7	296	2	T06572	Query Match		
23	73	20.7	613	2	S27770	Best Local Similarity 41.2%; Pred. No. 4.4e-05; DB 2; Length 566;		
24	73	20.7	919	2	A39248	Matches 28; Conservative 14; Mismatches 17; Indels 9; Gaps 4;		
25	72.5	20.5	185	2	S71512	QY 1 KRDPQREYEDRRRCQEQE-EPQYQCORC---RQORQHGRGCOLINPORGSSR		
26	71.5	20.3	679	2	B745262	Db 35 ERDPQ-QYQCORCQRESEATEERQEQCORCEREYKRCORQDE-ELQROYQOCGR		
27	70.5	20.3	554522			91		
28	71.5	20.3	910	2	A34721	RESULT 2		
29	71.5	20.3	911	2	B34721	S08059		

alpha-globulin type B precursor (tandem 1) - upland cotton (fragment)
N-Alternate names: seed storage protein
C-Species: *Cossypium hirsutum* (upland cotton)
C-Accession: S08059
C-Accession: S08059
R-Chian, C.A.; Borroto, K.; Kamaley, J.A.; Dure III, L.
Plant Mol. Biol. 9, 533-546, 1987
A-title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX.
A-Reference number: S06398
A-Accession: S08059
A-Status: not compared with conceptual translation

trichohyalin - hum
conserved hypothetical
hepatoma-derived g
protein kinase (EC
trypoblast-endoth
albumin 4 - ester
gag polypeptide -
troponin-I - scal
hypothetical prote
homocytic protein C
gag polypeptide -
transcription fact
hypothetical prote
hypothetical prote
arginine/glutamate
globulin-10 - maize

A; Molecule type: DNA
 A; Residues: 1-509 <CHL>
 C; Superfamily: glycinin

Query Match
 Best Local Similarity 33.4%; Score 118; DB 2; Length 509;
 Matches 34; Conservative 10; Mismatches 14; Indels 34; Gaps 5;

Qy 3 DPOQREYEDCRRRQEQQEPROQYQCCQRCA----REQQ-----RQH----- 38
 Db 1 DPQRR-YEECQEQCRRQEEQPCQPCQQRKRFQEQQOSOROFECQOQHCHOEQRERK 59

Qy 39 -----GRGDLINPQGGSGRVEEGERKQS 63
 Db 60 KQOCVRECREKYQENPWRG--EREEAEAEET 89

RESULT 3
 FWCNAB
 alpha-globulin B precursor (clone C72) - upland cotton
 N; Alternate names: seed storage protein; vicilin precursor
 C; Species: *Gossypium hirsutum* (upland cotton)
 C; Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C; Accession: A30838; S06911
 R; Chian, C.A.; Pyle, J.B.; Legocki, A.B.; Dure III, L.
 Plant Mol. Biol. 7, 475-489, 1986
 A; Title: Developmental biochemistry of cottonseed embryogenesis and germination XVIII. sc
 A; Reference number: A30838
 A; Accession: 1-588 <CHL>
 A; Cross-references: GB: M16891; NID: 9167374; PID: AAA33071.1; PID: 9167375
 A; Experimental source: var. Coker 201
 R; Chian, C.A.; Borroto, K.; Kamalay, J.A.; Dure III, L.
 Plant Mol. Biol. 9, 533-546, 1987
 A; Title: Developmental biochemistry of cottonseed embryogenesis and germination. XIX. sc
 A; Reference number: S06911
 A; Status: not compared with conceptual translation
 A; Molecule type: DNA
 A; Residues: 1-81 <CH2>
 C; Comment: This is a seed storage protein.
 C; Superfamily: glycinin
 C; Keywords: glycoprotein; seed; storage protein
 F; 1-25/domain: signal sequence #status predicted <SIG>
 F; 26-588/product: alpha-globulin storage protein #status predicted <SIG>
 F; 417/Binding site: carbohydrate (Asn) (covalent) #status predicted <SIG>

Query Match
 Best Local Similarity 31.2%; Score 110; DB 1; Length 588;
 Matches 33; Conservative 12; Mismatches 14; Indels 40; Gaps 6;

Qy 3 DPOQREYEDCRRRQEQQEPROQYQCCQRCA----REQQ-----RQHGRGDLINPQ 48
 Db 82 DPQRR-YEECQEQCRRQEEQPCQPCQQRKRFQEQQOSOROFECQOQHCHOEQRERK 139

Qy 49 RGGS-----GRY-----EEGEERKQS 63
 Db 140 KQOCVRECREKYQENPWREREEBETEEGEQEQS 178

RESULT 4
 S35221
 globulin Beg1 precursor - barley
 C; Accession: S35221
 R; Heck, G.; Chamberlain, A.K.; Ho, T.H.D.
 Mol. Gen. Genet. 239, 209-218, 1993
 A; Title: Barley embryo globulin 1 gene, Beg1: characterization of cDNA, chromosome mapping
 A; Reference number: S35221; MUID: 93287988

Query Match
 Best Local Similarity 33.3%; Score 101.5; DB 2; Length 605;
 Matches 20; Conservative 12; Mismatches 7; Indels 21; Gaps 2;

Qy 3 DPOQREYEDCRRRQEQQEPROQYQCCQRCA----REQQ-----RQHGRGDLINPQ 62
 Db 79 DPQRR-YEECQEQCRRQEEQPCQPCQQRKRFQEQQOSOROFECQOQHCHOEQRERK 117

RESULT 6
 T44130
 protein Pv100 (imported) - winter squash
 C; Species: Cucurbita maxima (winter squash)
 C; Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jul-2000
 C; Accession: T44130
 R; Yamada, K.; Shimada, T.; Kondo, M.; Nishimura, M.; Hara-Nishimura, I.
 J. Biol. Chem. 274, 2563-2570, 1999
 A; Title: Multiple functional proteins are produced by cleaving Asn-Gln bonds of a sin
 A; Status: preliminary; translated from GB/EMBL/IDB
 A; Accession: T44130
 A; Reference number: 222767; MUID: 99107919
 A; Molecule type: mRNA
 A; Residues: 1-810 <YAM>
 A; Cross-references: EMBL: AB019195; NID: 93808061; PID: BAN34056.1; PID: 93808062

Query Match
 Best Local Similarity 25.4%; Score 89.5; DB 2; Length 810;

